



Early Journal Content on JSTOR, Free to Anyone in the World

This article is one of nearly 500,000 scholarly works digitized and made freely available to everyone in the world by JSTOR.

Known as the Early Journal Content, this set of works include research articles, news, letters, and other writings published in more than 200 of the oldest leading academic journals. The works date from the mid-seventeenth to the early twentieth centuries.

We encourage people to read and share the Early Journal Content openly and to tell others that this resource exists. People may post this content online or redistribute in any way for non-commercial purposes.

Read more about Early Journal Content at <http://about.jstor.org/participate-jstor/individuals/early-journal-content>.

JSTOR is a digital library of academic journals, books, and primary source objects. JSTOR helps people discover, use, and build upon a wide range of content through a powerful research and teaching platform, and preserves this content for future generations. JSTOR is part of ITHAKA, a not-for-profit organization that also includes Ithaka S+R and Portico. For more information about JSTOR, please contact support@jstor.org.

Statistics and Its Methods

The History of Statistics. Their Development and Progress in Many Countries. In Memoirs to Commemorate the Seventy-Fifth Anniversary of the American Statistical Association. Collected and edited by JOHN KOREN. (New York: The Macmillan Company. 1918. Pp. xii, 773. \$7.50.)

This is, I believe, the only book in English, certainly the only important one, on the history of statistics, a remarkable fact when we remember that statistics as ordinarily conceived originated in London and that the oldest societies devoted to the subject are in England and the United States. A comparison between the present work and books or articles in other languages carrying the same or a similar title shows how differently the subject is conceived. Writings in French, Italian, or German on the history of statistics describe the development of statistics under the hands of private authors or university teachers; this book describes its development by legislators and administrators. The difference is suggested in the "their" of the title, for statistics as a collection of facts expressed in numerical form is a plural noun, but statistics as a method or a science, like mathematics, is a singular noun. The same difference is suggested by the introduction and by the title of the second and main division of the work, History and Development of Official Statistics in Many Countries. But its clearest statement is at the beginning of Professor Faure's memoir on the development of statistics in France. He says: "In the domain of history it has long been the habit . . . to confuse the practice of statistics with the theory of statistics. . . . It is necessary to study them separately . . . just as it is important to study separately the history of economic fact and the history of economic theory" (p. 217).

The present work, then, is mainly a history of the practice of statistics or of the efforts made by the various countries to gather and present information in numerical form about their own life and activities. Nearly one third of the space is devoted to the United States, with one address by the editor on the History of the Association, another by Dr. North on Seventy-five Years of Progress in American Statistics, and two memoirs, one on Federal Statistics, by Dr. John Cummings, the other on State Statistics, by Director Charles Gettemy.

In writing the history of statistical practice it is evidently hard to find and keep the thread of connection. The writers of these

memoirs do not fully agree among themselves as to what that thread is. Some evidently take the acts of the legislative and administrative agencies as their subject-matter; others regard the practice of statistics and its development as largely a product of individuals and give generous space for describing their work. Apparently the American writers center their attention on federal and state statutes, the European writers give more space to the persons involved. Perhaps the most significant contribution on American statistics is that of Dr. Cummings on Federal Statistics, which fills more than one seventh of the volume and more space than the three articles on other divisions of American statistics. In this valuable monograph the administrative statistical agency, like the Bureau of Education, is the unit for study. I doubt that this is better, though it certainly is easier, than to center attention on the kinds of statistical information obtained. Surely it would be more helpful to inform the inquirer that he could get some kinds of educational statistics from the annual reports of the Bureau of Education, illiteracy statistics from the federal censuses and those of certain states, etc., and to trace how these returns grew up and how the agencies gathering them have changed and developed. Perhaps it would have been still better to trace also the part that notable men have played in the history of American statistics, showing, for example, the influence of men like Shattuck, Walker, and Billings on our census development, Wright on the development of labor statistics, etc. A history of legislative and administrative changes tends to become mechanical; it is the persons behind these changes who vitalize the story.

In this respect the writers on foreign statistics seem as a rule to have interpreted their themes in a more significant way. Sir Athelstane Baines rightly puts John Graunt (not Grant) at the head of his tale of the growth of British statistics and finds fitting places for Halley and Sinclair, Rickman and Porter; Dr. Julin makes very clear the seminal influence of Quetelet in Belgium; Professor Faure in what seems to me the most valuable, as with the exception of that of Dr. Cummings it is the longest, paper in the volume shows how in their several fields Bodin and Sully and Colbert and Deparcieux and Necker and others contributed to build up the present system of French statistics.

But in pointing out this superiority of the best foreign work in the volume to that of our American scholars it would be unjust to

leave the impression that the latter are seriously at fault. Rather is it true that they have had to do pioneer work blazing trails or breaking ground while their European colleagues have had the work of many predecessors to build upon. Let me close by expressing my gratitude to the editor, his collaborators and the Association itself for this dignified and worthy memorial volume, and my belief that further monographic studies in this field are needed, some of which might profitably take as themes the statistical work and contributions of our best American statisticians.

WALTER F. WILLCOX.

Cornell University.

Statistical Methods Applied to Education. By HAROLD O. RUGG.
(Boston: Houghton Mifflin Company. 1917. Pp. xvii,
410. \$2.)

Another textbook in statistics has been added by Professor Rugg to our scant but growing list. In addition to chapters on the use, collection, tabulation, classification, and graphic presentation of educational data, there is one chapter on averages, one on variability, two on frequency curves, and one on correlation. The scope of the book, therefore, includes the practical phases of educational statistics and the simpler portions of statistical theory; and the treatment makes it an elementary text.

The material of the book is from the field of education. And the fact that the author is a professor of education raises the expectation of a superior teachable presentation of the subject. I think this natural expectation is fully met. My students who have used it as reference all testify that the presentation makes the methods easy to understand. The first chapter presents several most interesting results and uses of statistics without explaining the method or technique. This device greatly stimulates the interest of the student. The chapter on correlation is excellent from the pedagogical point of view. In this chapter the regression equations are discussed at length before the coefficient of correlation is introduced, which is, I think, the most teachable method. Many textbooks in mathematics and statistics are difficult for non-mathematical beginners because the explanations are greatly condensed. Professor Rugg's book, on the other hand, impresses one as sparing no words to explain.

The most noticeable defect is the treatment of the theory of

sampling. Sampling is discussed very scantily as a subdivision of the chapter on the normal frequency curve. Perhaps the author would reply that sampling is advanced statistical theory. Nevertheless, it is most important even for the beginner. Also the discussion of sampling not only seems inadequate but seems here and there to give the wrong impression. For instance, in discussing (p. 380) the fitting of the theoretical curve to the actual curve he suggests testing the frequency of the ordinates by subtracting or adding $3\sqrt{npq}$, saying nothing about the fact that although all ordinates may fall within these limits yet the actual curve will diverge from the normal curve beyond the limits of probability, because of the fact that it may be very improbable that two or more probable events will occur at the same time.

Furthermore, the author leaves the reader with the idea that the true arithmetic mean as contrasted with the sample arithmetic mean is based on all the cases that exist rather than on a theoretically infinite number of cases. "The *true* average spelling ability in the sixth grade of a large city system could be found by testing *all* of the 20,000 children, say, in all of the sixth grades in the system" (p. 224). If there were only five, say, Sanskrit scholars in the world, five would still be only a very small sample of a theoretically large number.

In discussing (p. 225) the distribution of averages of samples of 200 cases, he says "the best assumption" we can make is that they accord with the probability curve. This is a very important point in the theory of sampling. "The best assumption" could be considerably strengthened by reference to the extensive mathematics to prove that these averages follow the normal probability curve.

Also the designation of \sqrt{npq} as "the standard error of sampling" (p. 380 and p. 213) although commonly done, seems to be rather misleading description, as there are, of course, many other standard errors of samples.

The pitfalls of deceptive and spurious correlation are many. Professor Rugg warns of these, though sometimes vaguely (p. 254). ". . . such devices do not supply proofs of existing relationships . . . they are valuable in so far as they agree with sound logical analysis." The author also warns against the use of the correlation coefficient where the numbers are small, and has some very excellent discussion of this point. This is very important and timely, as a good many psychologists and education-

alists are using the coefficient of correlation when the number of cases are very few, sometimes as few as 10 or 12.

One might also quarrel with the author on his notation. Nearly every writer has some special notation peculiarly his own; that is to say, the formulas and constants in our text books have quite a variety of letters to represent each. Pedagogically, it would be better to approach still greater unity. The author uses in his chapter on variability the letter d to denote the variable differences between the arithmetic mean or the arbitrary origin of a variable and any number in the variable; while in the chapter on correlation the identical same difference is designated by x and x' (and y and y'). Σ which usually means "the sum of" is used in another place (p. 278) to equal $\sqrt{\frac{S(n_x(\bar{y}_x - \bar{y})^2)}{N}}$. There are other questionable points in notation.

A few errors in printing or writing are almost inevitable in mathematical books. The following were observed. On page 20 and page 275, $r = \frac{101.75}{154.6} = 65.8$; and on p. 229, $\sigma_{\text{diff between } x \text{ and } y} = \sqrt{\sigma_{m \text{ of } x}^2 - \sigma_{m \text{ of } y}^2}$. This should be $\sqrt{\sigma_{m \text{ of } x}^2 + \sigma_{m \text{ of } y}^2}$. Also, on page 252 $r = \frac{\Sigma xy}{N \sigma_x \sigma_y}$. The second σ should have the subscript y .

The book contains nothing on partial correlation. I have always found that beginners get the idea of partial correlation as easily as that of correlation and are quite fascinated with it. Often in social statistics, partial correlation makes otherwise worthless correlation coefficients full of meaning.

An excellent bibliography of educational statistics is included. The statistical bibliography is brief, but I think should certainly contain Pearson's *Tables for Biometricians and Statisticians* which came out in 1914.

WILLIAM F. OGBURN.

University of Washington.

NEW BOOKS

ALLER, C. C. *An industrial survey of Seattle*. (Seattle: Bureau of Industrial Research, University of Washington. 1918. Pp. 64.)

During the first part of 1917 the investigation here presented was conducted by Mr. Aller as research fellow of the University of Washington under the auspices of the Seattle Chamber of Commerce, the Commercial Club and Manufacturers Association of Seattle. Industrial growth is considered under markets, raw ma-

- terials, transportation, and manufacturing costs. As far as possible statistical and descriptive data are tabulated around these points.
- CLARK, W. W. *A statistical study of 102 truants*. Reprinted from the *Journal of Delinquency*, vol. 3, no. 5. (Whittier, Calif.: Whittier State School. 1918. Pp. 213-234.)
- FLORENCE, P. S. *Use of factory statistics in the investigation of industrial fatigue*. Columbia Univ. studies in political science, vol. LXXXI, no. 3. (New York: Longmans. 1918. Pp. 153. \$1.25.)
- HARRIS, L. I. *A clinical study of the frequency of lead, turpentine and benzin poisoning in four hundred painters*. Reprinted from the *Archives of Internal Medicine*, August, 1918. (Chicago: American Medical Association, 535 North Dearborn St. 1918. Pp. 28.)
- HARTLEY, E. F. *Census of the Virgin Islands of the United States, November 1, 1917*. (Washington: Bureau of the Census. 1918. Pp. 174.)
- HOFFMAN, F. L. *Army anthropometry and medical rejection statistics*. (Newark, N. J.: The author, Prudential Ins. Co. 1918.)
- KNIBBS, G. H. *Commonwealth and demography, 1917, and previous years*. Population and vital statistics bulletin no. 35. (Melbourne: Commonwealth Bureau of Census and Statistics. 1918. Pp. 285.)
- Colonies and dominions, from 1901 to 1915*. Statistical abstract, 53d number. (London: King. 1918. 2s.)
- Statistics of income. Compiled from the returns for 1916 under the direction of the Commissioner of Internal Revenue*. (Washington: Supt. Docs. 1918. Pp. 391.)
- Resumen anual de estadística municipal (año XV, 1917)*. (Montevideo: Dirección de Censo y Estadística. 1918. Pp. 346.)